Choosing the right price panel for attend home deliveries

André Morim
andre.morim@ltplabs.com
LTPlabs

David Viana
david.viana@ltplabs.com
LTPlabs

Fábio Neves-Moreira
fabio.s.moreira@inesctec.pt
INESC TEC

30/03/2025

Abstract: In the competitive landscape of Attended Home Deliveries in e-commerce, retailers must balance profitability with operational efficiency. Dynamic pricing strategies for delivery time slots are a key lever to achieve this balance, influencing both revenue generation and logistics costs. This white paper presents the application of the TRUST-AI[1] concept to online retail, addressing a dynamic Time Slot Pricing Vehicle Routing Problem. By modeling this challenge as a Markov Decision Process and solving it using symbolic expressions evolved through a Genetic Programming algorithm, we develop a learning framework that generates explainable and effective decision policies. Our approach goes beyond merely setting optimal delivery prices—it seeks to strategically nudge customers toward time slots that enhance route efficiency, ensuring multiple deliveries can be accommodated in the same trip. The goal is to maximize total profit, defined as the sum of revenue obtained from delivery fees and basket values of customers who proceed with their purchases, minus transportation costs, which include both fixed fleet costs and variable costs per kilometer. By integrating explainable AI into pricing and logistics decision-making, our methodology empowers retailers to dynamically adjust delivery fees while improving route consolidation, ultimately driving higher profitability and customer satisfaction.

1 Challenge

The rapid growth of online retail has intensified the complexity of efficiently managing the delivery of deliveries to the home. In this business sector, retailers set dynamic delivery fees that account for multiple factors, including customer preferences, demand fluctuations, and operational constraints

[1]TRUST-AI is a project funded by the European Union's Horizon 2020 research and innovation programme. More details available at `https://cordis.europa.eu/project/id/952060`.

such as fleet capacity (Agatz et al. 2011). In this challenge, our primary objective is to establish an optimized pricing strategy that maximizes total profit, comprising basket value and delivery fees, while minimizing logistics costs, comprising transportation and fleet costs.

The problem can be formulated as a Dynamic Time Slot Pricing Problem (DTSPP), where an online retailer must determine the best pricing strategy for each customer to maximize both revenue and logistics efficiency. Each customer, upon checkout, is presented with a set of available delivery slots, each priced differently based on demand. Customers may choose a slot or abandon the purchase if there is no suitable option.

This problem is inherently sequential and uncertain, requiring decisions to be made dynamically as new customers arrive. Moreover, on the retailer side, optimal time slot pricing must account for vehicle routing constraints, ensuring that deliveries are scheduled in a way that minimizes transportation costs while maintaining a high service level, profiting from an available fixed fleet of capacitated vehicles. The challenge extends beyond simply assigning prices, it involves influencing customer behavior to create delivery routes that reduce operational costs by consolidating orders efficiently (Klein et al. 2018). A representation of this sequential decision problem is presented in Figure 1.

Traditional pricing models often overlook the logistical impact of customer selections, leading to inefficiencies in routing and increased delivery costs. By considering customer choice probabilities and vehicle routing constraints, a more intelligent pricing strategy can be deployed, ensuring that profitable and logistically efficient decisions are made throughout the booking horizon.

This problem can be addressed by an AI-driven approach capable of learning from past decisions and dynamically adapting to new customer arrivals, ensuring that pricing strategies align with both customer preferences and logistical efficiency.

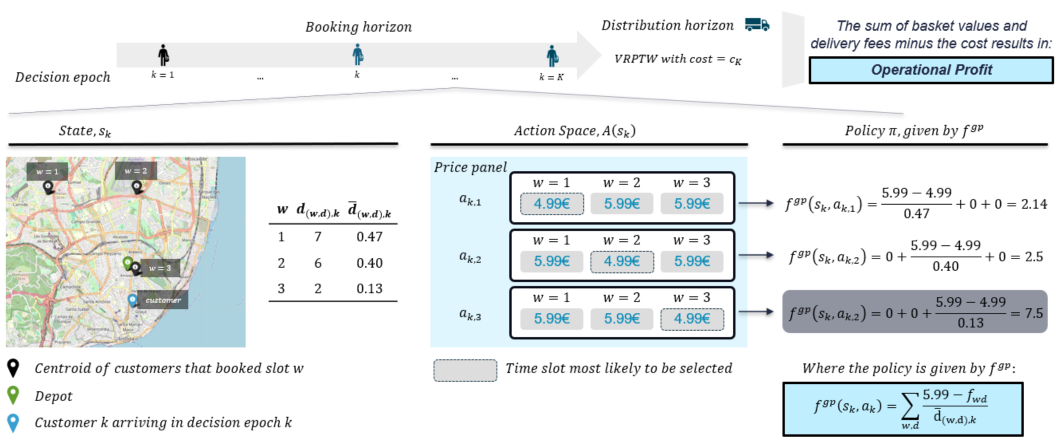


Figure 1: Overview of the DTSPP and its methodological components.

2 Solution

To address this challenge, we employ the TRUST-AI concept, which consists on a guided-empirical learning process to obtain explainable symbolic AI models.

Our approach to solving the dynamic time slot pricing problem integrates machine learning techniques with logistics optimization to create an intelligent, explainable, and adaptable pricing model. By leveraging a Markov Decision Process (MDP) formulation within the OpenAI Gym framework, we ensure that decisions are made sequentially, adapting to real-time customer behavior and logistics constraints. Unlike conventional pricing methods, our approach utilizes symbolic expressions to determine optimal pricing actions, which allows for greater transparency and adaptability.

A key innovation of this methodology is its ability to develop integrated pricing policies that consider both customer preferences and logistical efficiency. Traditional prescriptive pricing models often neglect the operational impact of customer choices, resulting in inefficient route planning and increased delivery costs. Our approach directly addresses this issue by balancing revenue maximization with fleet and cost optimization.

Furthermore, this framework is designed to involve operations managers in the decision-making process. By defining terminal and mathematical operators for symbolic expressions, managers can fine-tune policies to align with business objectives. This unique aspect enables businesses to maintain control over pricing complexity, ensuring that AI-generated policies remain actionable and interpretable (Adadi and Berrada 2018; Murdoch et al. 2019).

2.1 Problem Formulation as an MDP

Our approach is built on an MDP formulation within the OpenAI Gym framework, enabling a structured environment for sequential decision-making under uncertainty. This formulation provides a controlled simulation setting to evaluate different pricing policies and their impact on both customer behavior and logistics constraints.

The key components of our MDP are as follows. **State Space:** Represents the logistics system conditions, including current order distribution, fleet availability, and time-dependent demand trends. **Action Space:** Defines the possible pricing panels to be presented to customers, influencing their selection of time slots. **Transition Function:** Captures probabilistic customer responses to pricing, incorporating customer willingness-to-pay models and historical demand patterns. **Reward Function:** Quantifies profitability including immediate revenue from fees and basket values against longer-term cost considerations such as delivery efficiency and fleet utilization.

This formulation enables a pricing strategy to be evaluated, which is a key component of every policy learning framework.

2.2 Policy Learning via Genetic Programming

A key innovation in our method is the use of symbolic expressions to score different price panels (actions). Unlike traditional prescriptive models, which often rely on black-box machine learning techniques, our approach enables the generation of interpretable policies that can be directly inspected and adjusted by operations managers.

To derive symbolic expressions that serve as decision policies for price panel selection, we use Genetic Programming (GP) (Koza 1994). The learning process iteratively refines symbolic expressions to maximize cumulative rewards, learning which strategies yield optimal profitability and efficient routing outcomes. An example of a decision policy is represented by mathematical expression (1).

$$(p_1 - oi) + \frac{1}{n} \cdot \left(\frac{oi \cdot p_1}{n^2} - oi + \overline{p} - \frac{\frac{oi}{n}}{n - oi} \right) \tag{1}$$

The four features used in this sample decision policy include: p_1, the delivery fee of the preferred customer time slot; oi, a measure of occupation imbalance between all time slots subject to planning (i.e., the higher the volume of orders of a particular time slot compared to others, leads to higher values of oi); n, represents the number of customers who already booked a time slot; and $\overline{p}$, the average delivery fee of the panel presented to the customers.

In the context of the DTSPP, the candidate price panel that maximizes expression (1) is the one presented to the customer. Therefore, this decision policy previliges increasing the delivery fee of the customer's preferred time slot, while trying to minimize the imbalance in the selection of time slots, measured by feature oi. GP also learns a term concerned with increasing the overall delivery fee (through feature $\overline{p}$) that decreases its importance the number of customers already booked increases (through term $1/n$).

This approach entails three key aspects. **Explainability-by-Design:** Symbolic expressions provide human-readable pricing policies, allowing operations managers to understand and refine decision-making rules. **Integrated Pricing Policy:** The learned policies incorporate both customer preferences and logistics constraints, ensuring that pricing not only maximizes revenue but also enhances route efficiency. **Adaptive Decision Logic:** The learning framework is capable of incorporating and adapting to different operational manager preferences, as well as different logistics contexts (e.g., account for different customer concentration).

This rare application of symbolic expressions in prescriptive problems allows for greater transparency in AI-driven decision-making, bridging the gap between automated optimization and business expertise.

2.3 Integration into TRUST Platform

The proposed learning framework is implemented within the TRUST platform[2], leveraging customizable dashboards for training, testing and visualizing symbolic policies. The interactive framework supports human-algorithm collaboration, enabling iterative refinement and adaptation to business needs. The platform allows for scenario analysis, comparing different pricing policies to identify trade-offs between revenue maximization and cost minimization, as well as symbolic expression size and complexity. Figure 2 provides a look at TRUST's platform being parametrized to derive pricing policies for the DTSPP.

The generated runs can then be inspected through visualizations. Figure 3 provides two visualizations plotted for a given training session. Figure 3 (a) provides an approximation front contrasting model performance against model complexity. Alternatively, Figure 3 (b) displays the approximation front comparing the revenue accrued by each model against the resulting variable cost.

[2]For more information on this Symbolic AI platform, the reader is referred to `https://gitlab.inesctec.pt/trust-ai/framework`

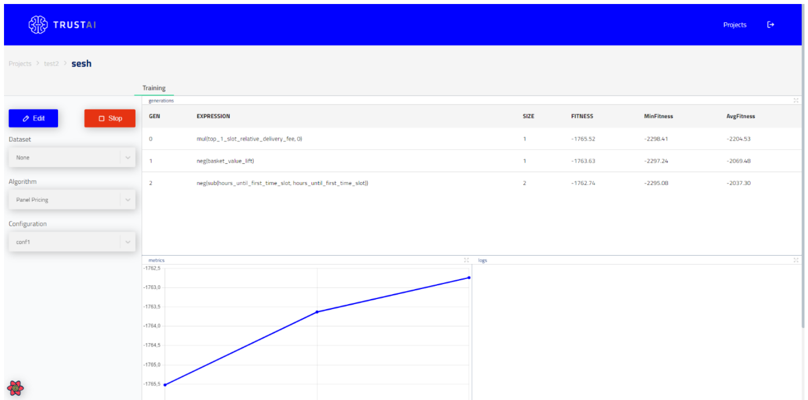
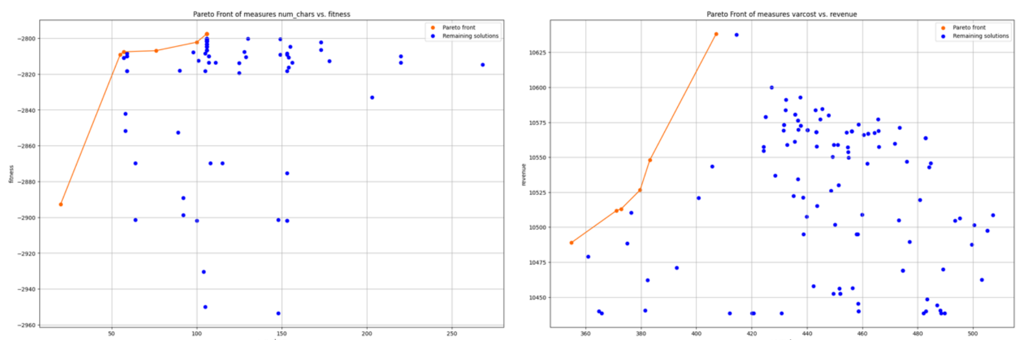


Figure 2: A sample run of TRUST platform being used to derive decision policies for the DTSPP.



(a) (b)

Figure 3: Examples of visualizations provided by the TRUST platform to inspect learned AI models.

3 Results

To validate the proposed learning framework in an online retail setting, we conducted three validation workshops in collaboration with the retailer who inspired this work. These workshops allowed us to assess the effectiveness of our approach in a real-world context, gathering feedback from industry professionals to refine and optimize both the realism of the simulation environment and the pricing policies. The sessions focused on evaluating the explainability and adaptability of the symbolic expressions, ensuring that they aligned with both business objectives and operational constraints. Several important achievements can be emphasized from this relationship including a research institute, a consultancy company, and a retailer:

Profit Maximization: The learned symbolic expressions effectively balance revenue generation (basket value + delivery fees) with transportation cost minimization by encouraging customer selections that optimize route consolidation.

Explainability and Interpretability: Unlike black-box AI models, the GP-based approach provides human-readable policies that allow business stakeholders to understand, trust, and refine pricing strategies.

Efficient Resource Utilization: The dynamic pricing model improves vehicle routing efficiency by influencing customer selections toward cost-effective delivery slots, reducing fleet operational costs and optimizing last-mile logistics.

Adaptability to Business Constraints: The methodology can be customized to incorporate retailer-specific objectives, such as sustainability considerations (e.g., encouraging off-peak delivery windows) or priority customer handling.

Scalability and Robustness: The model is designed to adapt to fluctuations in customer demand, enabling retailers to dynamically optimize pricing strategies as operational constraints evolve.

Beyond discussing the TRUST-AI concept, the DTSPP simulator, and explainability concerns (e.g., whether an operations manager could devise a decision policy based on domain knowledge), we validated our methodology by applying it to a toy problem instance and analyzing the results. Figure 4 compares the results obtained using three distinct decision policies:

Tactical (As Is): Static approach where the retailer sets time slot prices periodically based on expected occupancy to achieve a break-even situation, without adjusting prices for each customer arrival.

Myopic: Treedy dynamic approach in which, at each customer arrival, a time slot price panel that maximizes the expected revenue is presented, while disregarding subsequent impact on distribution costs.

Dynamic: Decision policy represented by expression (1) learned using GP to balance revenue maximization with operational efficiency.

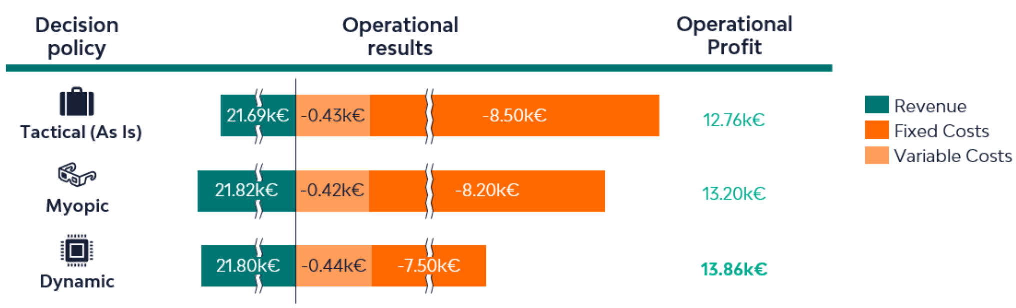


Figure 4: Comparison of three decision policies applied to a toy instance of the DTSPP.

As illustrated in Figure 4, transitioning from a static decision policy to dynamic approaches results in an expected increase in operational profit of 3.4% and 8.6% for the myopic and dynamic strategies, respectively. The reduction in fixed costs observed in the myopic approach, relative to the tactical strategy, can be attributed to customers being incentivized to select their preferred time slots, thereby enhancing delivery consolidation within the same vehicles.

Furthermore, a comparison between the myopic and dynamic approaches confirms that our GP-based methodology effectively learned the opportunity cost associated with each pricing decision. Notably, while incurring only a 0.09% reduction in revenue and achieving an 8.54% decrease in fixed costs, our decision policy successfully maximized revenue while accounting for the impact of each pricing decision on overall distribution costs. Ultimately, when benchmarked against the retailer's current strategy, our approach achieved an 11.76% reduction in fixed costs.

4 Conclusion

The application of TRUST-AI in online retail showcases the potential of explainable AI in solving complex prescriptive problems. By leveraging symbolic expressions for dynamic time slot pricing, retailers can enhance profitability while maintaining operational efficiency. This approach provides a transparent, adaptable, and scalable solution to the challenges of last-mile delivery optimization, bridging the gap between pricing strategy and logistics execution.

Acknowledgments

The research leading to these results has received funding from the European Union's Horizon 2020 - The EU Framework Programme for Research and Innovation 2014-2020, under grant agreement 952060.

References

Adadi, Amina and Mohammed Berrada (2018). "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)". In: *IEEE Access* 6, pp. 52138–52160. DOI: 10.1109/ACCESS.2018.2870052.

Agatz, Niels et al. (2011). "Time Slot Management in Attended Home Delivery". In: *Transportation Science* 45.3, pp. 435–449. DOI: 10.1287/trsc.1100.0346.

Klein, Robert et al. (Oct. 2018). "A Model-Based Approximation of Opportunity Cost for Dynamic Pricing in Attended Home Delivery". In: *OR Spectr.* 40.4, pp. 969–996. ISSN: 0171-6468. DOI: 10.1007/s00291-017-0501-3. URL: https://doi.org/10.1007/s00291-017-0501-3.

Koza, John R. (1994). "Genetic programming as a means for programming computers by natural selection". In: *Statistics and Computing* 4.2, pp. 87–112. DOI: 10.1007/BF00175355.

Murdoch, W. James et al. (Oct. 2019). "Definitions, methods, and applications in interpretable machine learning". In: *Proceedings of the National Academy of Sciences* 116.44, pp. 22071–22080. DOI: 10.1073/pnas.1900654116. URL: https://doi.org/10.1073/pnas.1900654116.